

Sonderdruck aus:

Mitteilungen aus der Arbeitsmarkt- und Berufsforschung

Horst-Manfred Schellhaaß und Alexandra Schubert

Internationale Entwicklungen der
Evaluierungsmethoden für arbeitsmarktpolitische
Programme

25. Jg./1992

3

Mitteilungen aus der Arbeitsmarkt- und Berufsforschung (MittAB)

Die MittAB verstehen sich als Forum der Arbeitsmarkt- und Berufsforschung. Es werden Arbeiten aus all den Wissenschaftsdisziplinen veröffentlicht, die sich mit den Themen Arbeit, Arbeitsmarkt, Beruf und Qualifikation befassen. Die Veröffentlichungen in dieser Zeitschrift sollen methodisch, theoretisch und insbesondere auch empirisch zum Erkenntnisgewinn sowie zur Beratung von Öffentlichkeit und Politik beitragen. Etwa einmal jährlich erscheint ein „Schwerpunktheft“, bei dem Herausgeber und Redaktion zu einem ausgewählten Themenbereich gezielt Beiträge akquirieren.

Hinweise für Autorinnen und Autoren

Das Manuskript ist in dreifacher Ausfertigung an die federführende Herausgeberin Frau Prof. Jutta Allmendinger, Ph. D.
Institut für Arbeitsmarkt- und Berufsforschung
90478 Nürnberg, Regensburger Straße 104
zu senden.

Die Manuskripte können in deutscher oder englischer Sprache eingereicht werden, sie werden durch mindestens zwei Referees begutachtet und dürfen nicht bereits an anderer Stelle veröffentlicht oder zur Veröffentlichung vorgesehen sein.

Autorenhinweise und Angaben zur formalen Gestaltung der Manuskripte können im Internet abgerufen werden unter http://doku.iab.de/mittab/hinweise_mittab.pdf. Im IAB kann ein entsprechendes Merkblatt angefordert werden (Tel.: 09 11/1 79 30 23, Fax: 09 11/1 79 59 99; E-Mail: ursula.wagner@iab.de).

Herausgeber

Jutta Allmendinger, Ph. D., Direktorin des IAB, Professorin für Soziologie, München (federführende Herausgeberin)
Dr. Friedrich Buttler, Professor, International Labour Office, Regionaldirektor für Europa und Zentralasien, Genf, ehem. Direktor des IAB
Dr. Wolfgang Franz, Professor für Volkswirtschaftslehre, Mannheim
Dr. Knut Gerlach, Professor für Politische Wirtschaftslehre und Arbeitsökonomie, Hannover
Florian Gerster, Vorstandsvorsitzender der Bundesanstalt für Arbeit
Dr. Christof Helberger, Professor für Volkswirtschaftslehre, TU Berlin
Dr. Reinhard Hujer, Professor für Statistik und Ökonometrie (Empirische Wirtschaftsforschung), Frankfurt/M.
Dr. Gerhard Kleinhenz, Professor für Volkswirtschaftslehre, Passau
Bernhard Jagoda, Präsident a.D. der Bundesanstalt für Arbeit
Dr. Dieter Sadowski, Professor für Betriebswirtschaftslehre, Trier

Begründer und frühere Mitherausgeber

Prof. Dr. Dieter Mertens, Prof. Dr. Dr. h.c. mult. Karl Martin Bolte, Dr. Hans Büttner, Prof. Dr. Dr. Theodor Ellinger, Heinrich Franke, Prof. Dr. Harald Gerfin,
Prof. Dr. Hans Kettner, Prof. Dr. Karl-August Schäffer, Dr. h.c. Josef Stingl

Redaktion

Ulrike Kress, Gerd Peters, Ursula Wagner, in: Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB),
90478 Nürnberg, Regensburger Str. 104, Telefon (09 11) 1 79 30 19, E-Mail: ulrike.kress@iab.de; (09 11) 1 79 30 16,
E-Mail: gerd.peters@iab.de; (09 11) 1 79 30 23, E-Mail: ursula.wagner@iab.de; Telefax (09 11) 1 79 59 99.

Rechte

Nachdruck, auch auszugsweise, nur mit Genehmigung der Redaktion und unter genauer Quellenangabe gestattet. Es ist ohne ausdrückliche Genehmigung des Verlages nicht gestattet, fotografische Vervielfältigungen, Mikrofilme, Mikrofotos u.ä. von den Zeitschriftenheften, von einzelnen Beiträgen oder von Teilen daraus herzustellen.

Herstellung

Satz und Druck: Tümmels Buchdruckerei und Verlag GmbH, Gundelfinger Straße 20, 90451 Nürnberg

Verlag

W. Kohlhammer GmbH, Postanschrift: 70549 Stuttgart; Lieferanschrift: Heßbrühlstraße 69, 70565 Stuttgart; Telefon 07 11/78 63-0;
Telefax 07 11/78 63-84 30; E-Mail: waltraud.metzger@kohlhammer.de, Postscheckkonto Stuttgart 163 30.
Girokonto Städtische Girokasse Stuttgart 2 022 309.
ISSN 0340-3254

Bezugsbedingungen

Die „Mitteilungen aus der Arbeitsmarkt- und Berufsforschung“ erscheinen viermal jährlich. Bezugspreis: Jahresabonnement 52,- € inklusive Versandkosten; Einzelheft 14,- € zuzüglich Versandkosten. Für Studenten, Wehr- und Ersatzdienstleistende wird der Preis um 20 % ermäßigt. Bestellungen durch den Buchhandel oder direkt beim Verlag. Abbestellungen sind nur bis 3 Monate vor Jahresende möglich.

Zitierweise:

MittAB = „Mitteilungen aus der Arbeitsmarkt- und Berufsforschung“ (ab 1970)
Mitt(IAB) = „Mitteilungen“ (1968 und 1969)
In den Jahren 1968 und 1969 erschienen die „Mitteilungen aus der Arbeitsmarkt- und Berufsforschung“ unter dem Titel „Mitteilungen“, herausgegeben vom Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit.

Internet: <http://www.iab.de>

Internationale Entwicklungen der Evaluierungsmethoden für arbeitsmarktpolitische Programme

*Horst-Manfred Schellhaaß und Alexandra Schubert, Berlin**

Ziel dieses Beitrages ist es, internationale Weiterentwicklungen der Evaluationsmethoden für die Erfolgskontrolle deutscher Arbeitsmarktpolitik nutzbar zu machen. Dabei betrachten wir drei Aspekte der Methodenentwicklung:

- 1) Bestimmung und Operationalisierung der Zielkriterien,
- 2) Evaluationsmethoden zur Ermittlung marginaler Kosten und Erträge eines Programmes,
- 3) Evaluationsmethoden, die neben dem zentralisierbaren auch das verstreut vorhandene Wissen der einzelnen Arbeitsmarktakteure nutzbar machen.

zu 1): Häufig wird in Evaluationen die Zielfunktion so bestimmt, daß sie die speziellen Wirkungen des untersuchten Programmes gut erfaßt. Wichtiger ist es jedoch, Zielkriterien zu verwenden, anhand derer verschiedene Maßnahmen zur Lösung eines arbeitsmarktpolitischen Problems miteinander verglichen werden können. Zur Beurteilung von Maßnahmen zur Eingliederung von strukturell benachteiligten Arbeitslosen ist insbesondere die Entwicklung des Einkommens geeignet. Im Maßnahmevergleich sollte dabei auf ähnliche Spezifikation der Zielfunktion und unterschiedlich hohe Maßnahmekosten geachtet werden.

zu 2): Methodisch sollten Wirkungsanalysen so konzipiert sein, daß die tatsächlich auf das Programm zurückzuführenden Veränderungen der Zielgrößen möglichst genau festzustellen sind. Traditionelle Evaluationen messen dabei nur die durchschnittlichen Erträge eines Programms. Wir zeigen, daß es inzwischen methodisch möglich ist, die marginalen Kosten und Erträge unterschiedlicher Teilnehmergruppen festzustellen. Mit Hilfe der Informationen, die durch diese Anpassung an die mikroökonomischen Optimierungsmodelle gewonnen werden, kann der Umfang eines Programms effizient gesteuert werden.

zu 3): Traditionelle Wirkungsanalysen nutzen nur das zentralisierbare Wissen. Sie können nicht feststellen, inwieweit das kurzlebige und spezialisierte Wissen über Marktgegebenheiten, das auf die handelnden Individuen verstreut ist, im Sinne einer Steigerung der gesamtwirtschaftlichen Effizienz zum Tragen kommt. Eine stärkere Betonung der Prozeßanalyse ermöglicht nach unserer Ansicht eine institutionelle Ausgestaltung der Arbeitsmarktpolitik, in der die zur Nutzung dezentralen Wissens notwendigen Spielräume und Anreize vorhanden sind.

Gliederung

- 1 Problemstellung
 - 2 Auswahl und Operationalisierung der Zielkriterien
 - 3 Durchschnitts- versus Marginalanalyse in Evaluationen
 - 4 Von der Wirkungsanalyse zur Prozeßanalyse
 - 5 Schluß
- Literaturverzeichnis

I Problemstellung

In der Bundesrepublik Deutschland wird aktive Arbeitsmarktpolitik im Prinzip von allen gesellschaftlichen Kräften befürwortet. Harte Kontroversen gibt es jedoch über Umfang, Ausgestaltung und Anwendungsbereich der einzelnen Instrumente. Evaluationen sollen zur Lösung der strittigen Punkte beitragen, indem sie den Erfolg oder Mißerfolg vergangener Projekte messen und Verbesserungsvorschläge für zukünftige Programme unterbreiten. Allerdings befinden sich die Evaluierungstechniken selbst

noch in ihrer Entwicklungsphase, so daß sich noch keine standardisierten Verfahren durchgesetzt haben. Das Ziel dieses Beitrages besteht darin, internationale Weiterentwicklungen der Evaluationsmethoden für die Erfolgskontrolle deutscher Arbeitsmarktpolitik nutzbar zu machen.

Unser erster Ansatzpunkt ist der Maßstab, an dem der Erfolg von Arbeitsmarktmaßnahmen gemessen werden soll. Hier ist die oft nicht explizit diskutierte Frage zu klären, welche Ziele bei der Evaluation zu berücksichtigen sind und wie ihre Operationalisierung erfolgen kann.

Nach der Diskussion zweckmäßiger Zielkriterien greifen wir als zweiten Ansatzpunkt die ökonomische Aussagefähigkeit der Ergebnisse von Wirkungsanalysen („impact analysis“) heraus. Über den gegenwärtigen Methodenstreit, ob für verlässliche Ergebnisse eine Zufallsauswahl der Teilnehmer notwendig ist, ist vernachlässigt worden, daß der Ausweis durchschnittlicher Programmwirkungen nur unzureichende Informationen für eine optimale Steuerung der aktiven Arbeitsmarktpolitik liefert. Inzwischen gibt es erste Ansätze, das für eine Optimierung erforderliche Marginalprinzip in die Evaluierungstechnik zu integrieren.

Unser dritter Ansatzpunkt setzt an der Frage an, ob traditionelle Wirkungsanalysen das gesamte relevante Wissen erfassen. Unsere These lautet: Wirkungsanalysen liefern die lange Zeit gültigen und deshalb zentralisierbaren Informationen, während die Prozeßanalyse („process analysis“)

* Prof. Dr. Horst-Manfred Schellhaaß ist Inhaber des Lehrstuhls für Mikroökonomie an der Technischen Universität Berlin; Diplom-Volkswirtin Alexandra Schubert ist wiss. Mitarbeiterin am gleichen Lehrstuhl. Der Beitrag liegt in der alleinigen Verantwortung der Autoren.

die institutionellen Rahmenbedingungen daraufhin überprüft, ob sie eine Nutzung des unvorhersehbaren kurzlebigen Wissens „vor Ort“ ermöglichen. Mit dieser Aufgabenzuweisung an Prozeß- und Wirkungsanalyse betreten wir wissenschaftliches Neuland.

2 Auswahl und Operationalisierung der Zielkriterien

Objektbereich der hier untersuchten Evaluationen ist die aktive Arbeitsmarktpolitik. Arbeitsmarktpolitik umfaßt nach unserer Vorstellung all jene Maßnahmen zur Beeinflussung von Niveau und Struktur der Beschäftigung, die am einzelnen Individuum ansetzen, im Gegensatz zur auf makroökonomischer Ebene wirkenden Konjunktur- oder Lohnpolitik. Aktiv wird diese Politik genannt, wenn sie arbeitsmarktrelevante Qualitäten der Individuen verändert. Nicht beachtet werden von uns daher Transferzahlungen, die die Einkommensverluste der Arbeitslosen finanziell kompensieren sollen, und Maßnahmen, deren Ziel die Ausgliederung aus dem Arbeitsprozeß ist (z. B. Vorruhestandsregelungen).

Aufgabe aktiver Arbeitsmarktpolitik ist es, Chancengleichheit auf dem Arbeitsmarkt für die temporär oder dauerhaft wettbewerbsschwachen Mitglieder der Gesellschaft herzustellen.¹ Die geminderte Wettbewerbsfähigkeit kann verschiedene Ursachen haben: zum Beispiel Immobilität, eine ungenügende oder nicht mehr verwendbare Qualifikation, eine aufgrund von gesundheitlichen Einschränkungen oder längerer Arbeitslosigkeit real oder im Vorurteil der Arbeitgeber geminderte Wertschöpfung.

Diese Aufgabe wird durch Arbeitsmarktprogramme zur Wiedereingliederung erfüllt, die in Europa meist als Rechtsanspruch der Betroffenen betrachtet werden. In den Vereinigten Staaten hingegen gelten sie als Investitionen, die für alle Beteiligten eine meßbare Rendite erbringen sollten (vgl. Gueron 1991, S. 167 f.). Folglich hat sich in den amerikanischen Evaluationen Effizienz als Erfolgsmaßstab weitgehend durchgesetzt. Zudem hat sich gezeigt, daß eine Herausstellung verteilungspolitisch motivierter Ziele kontraproduktiv im Hinblick auf die allgemeine Unterstützung aktiver Arbeitsmarktpolitik wirkt. So vermuten Havemann/Saks (1985), daß die Popularität des CETA-Programms² in dem Moment sank, als eine größere Konzentration auf besonders benachteiligte Gruppen erfolgte – obwohl die nachfolgenden Evaluationsergebnisse Hinweise darauf gaben, daß CETA für diese Gruppen besonders effizient war.

Auch wenn Effizienz als Maßstab akzeptiert ist, sind verschiedene Operationalisierungen dieses Zieles möglich. Wenige Analysen versuchen so viele denkbare Wirkungen abzudecken, wie die Evaluation des Job Development Programs in Kanada, in dem 24 Gruppen von Erfolgsfaktoren zur Beurteilung des Programmes herangezogen worden sind (Goss, Gilroy & Associates 1989, S. 8 ff.). Der Vorteil eines umfassenden Zielkriteriensystems ist, daß nicht nur die Auswirkungen auf Niveau und Struktur der Beschäftigten, sondern auch externe Effekte und Nebenwirkungen

erfaßt werden. Als Beispiel ziehen wir die Evaluation des Job-Corps-Programmes durch Long/Mallar/Thornton (1981) heran. Zielgruppe dieses Programmes waren Schulabbrecher, die in festen Zentren untergebracht wurden. Ihre Betreuung umfaßte ein breites Spektrum an Aktivitäten zur allgemeinen und beruflichen Bildung sowie zur gesundheitlichen Betreuung. Ziel war neben der Einkommenssteigerung auch eine verminderte Abhängigkeit von Ausgleichszahlungen und die Vermeidung unsozialen Verhaltens (insbesondere Kriminalität).

Ein von den Autoren gut gelöstes Problem ist die Unterscheidung von realen Veränderungen des volkswirtschaftlichen Ressourcenbestandes, die den Nettonutzen unmittelbar beeinflussen, auf der einen Seite und von Ressourcentransfers, die erwünschte oder unerwünschte Verteilungswirkungen aufzeigen, auf der anderen Seite. Intangible Größen werden qualitativ beschrieben. Ein Beispiel: Aus Sicht der Steuerzahler führt verminderte Diebstahlskriminalität dazu, daß ihre Wohlfahrt in dem Maße steigt, wie der Umfang gestohlener Güter – gemessen an deren Marktwert – abnimmt. Ein Teil dieses Wertes ist jedoch ein Verlust für die Programmteilnehmer, da er für sie entgangenes Einkommen darstellt. Der tatsächliche Zugewinn der Gesellschaft ist im geringeren Ressourcenverbrauch für Schutzeinrichtungen sowie darin zu sehen, daß weniger Gegenstände im Verlauf von Einbrüchen beschädigt werden. Die verminderte psychologische Belastung stellt einen intangiblen Nutzen auf beiden Seiten dar und wird als immaterieller Wohlfahrtsgewinn der Gesellschaft aufgefaßt (Long/Mallar/Thornton 1981, S. 61).

Häufig läßt sich das Zielkriteriensystem von Evaluationen deshalb vereinfachen, weil die Nebenwirkungen mit dem arbeitsmarktlichen Erfolg der Maßnahme korrelieren. Dies gilt ebenso für die abnehmende Kriminalität wie für die verminderte psychische Belastung der Arbeitslosen oder für den verhinderten Abbau von Humankapital als Folge von Langzeitarbeitslosigkeit. Insofern reicht im Vergleich verschiedener Maßnahmen in der Regel eine Konzentration auf die engen arbeitsmarktlichen Ziele aus.

Zur Beurteilung des arbeitsmarktlichen Erfolges ist der Einfluß des Programmes auf Beschäftigungsniveau und -Struktur zu schätzen. Da unmittelbar nur die Bruttowirkungen des Programmes gemessen werden können, müssen die Prozesse berücksichtigt werden, die zu einer Differenz zwischen der Bruttobeschäftigungssteigerung und dem Nettobeschäftigungsgewinn führen (vgl. Schellhaaß 1991, S. 94 f.). Dabei kommen drei Effekte in Betracht:

Der *Mitnahmeeffekt* kennzeichnet die Tatsache, daß einige Programmteilnehmer sich ohne das Programm in genau der gleichen Weise verhalten hätten. Das Problem ist also die Abschätzung der Status-quo-Entwicklung für die Teilnehmer, das meist durch die Bildung einer geeigneten Kontrollgruppe gelöst wird.

Mitnahmeeffekte können auch auftreten, wenn Unternehmen oder öffentliche Arbeitgeber durch Subventionen zur Beteiligung an Arbeitsmarktmaßnahmen ermuntert werden. Ein Beispiel sind öffentliche Arbeitsbeschaffungsprogramme, bei denen durch eine „Zusätzlichkeitsbedingung“ verhindert werden soll, daß reguläre Haushaltsmittel durch Programmgelder ersetzt werden („fiskalische Substitution“). In ihrer Untersuchung der Wirkung von PEP³ – und CETA-Maßnahmen zur Arbeitsbeschaffung kommen Johnson/Tomola (1977, S. 16 f.) zu dem Ergebnis, daß der Mitnahmeeffekt innerhalb des ersten halben Jahres nach

¹ Vorwiegend antizyklische Maßnahmen (z. B. Kurzarbeitergeld) sowie die vorbeugende Qualifizierung der Beschäftigten zur Verhinderung von gesamtwirtschaftlichen Qualifikationsdefiziten, werden hier aufgrund anders gelagerter Evaluierungsprobleme nicht explizit behandelt.

² „Comprehensive Employment and Training Act“. Das von 1974 bis 1981 laufende Programm umfaßte verschiedene Maßnahmen wie Arbeitsbeschaffung, on-the-job-Training, theoretische Ausbildung etc.

³ „Public Employment Program“, ein nur 1972 und 1973 laufendes Programm zur antizyklischen Arbeitsbeschaffung.

Einführung eines Programmes gering ist, ein Jahr später jedoch fast hundert Prozent beträgt⁴. Dies weist darauf hin, daß mit zunehmender Programmlaufzeit die Überprüfung der „Zusätzlichkeit“ immer schwieriger wird, da die antragstellenden Institutionen die Programme bereits in ihrer Budgetplanung berücksichtigen. Auch wenn zum Antragszeitpunkt keine Mittel für einen bestimmten Zweck zur Verfügung stehen, ist nicht auszuschließen, daß ohne die Programme eine andere Aufteilung der vorhandenen Gelder erfolgt wäre.

Ein zweiter Einflußfaktor auf die Differenz von Brutto- und Nettowirkungen sind *Substitutionseffekte*. Diese Effekte treten auf, wenn Programmteilnehmer nach Maßnahmen mit anderen Arbeitskräften um die gleichen Arbeitsplätze konkurrieren. Nur bei einer vollständig elastischen Arbeitsnachfrage oder der Qualifikation für einen Mangelberuf wäre die Beschäftigung und der Einkommenszuwachs der Teilnehmer nach Programmende tatsächlich zusätzlich. Wenn die Beschäftigung von Programmteilnehmern zu einer Verdrängung anderer Arbeitskräfte führt, kommt es zwar zu keiner Steigerung des Volkseinkommens, aber es ist ein Beweis dafür, daß die Bemühungen um die Herstellung der Wettbewerbsfähigkeit der Programmteilnehmer am Arbeitsmarkt erfolgreich waren. Wenn nicht durch gezielte Qualifikation der Programmteilnehmer das zusätzliche Arbeitsangebot auf einem Teilmarkt mit überschüssiger Nachfrage auftritt, sind bei der Berechnung des gesamtgesellschaftlichen Nutzens die Einkommensgewinne der Teilnehmer um die zusätzlichen Mobilitätskosten der substituierten Nichtteilnehmer zu kürzen.

Als dritter Einflußfaktor sind *Verdrängungseffekte* zu berücksichtigen. Diese spielen bei der Bewertung des von den Programmteilnehmern im Laufe des Programms erstellten Outputs eine Rolle. Long/Mallar/Thornton (1981, S. 60 f.) bewerten den Output eines Teilnehmerarbeitstages mit dem durchschnittlichen Preis, den andere Anbieter für die Bereitstellung des Outputs gefordert hätten, abzüglich der direkten Kosten der Outputerzeugung. Dieser Ansatz ist so lange sinnvoll, wie das zusätzliche Angebot der mit Hilfe der Fördermaßnahme erzeugten Güter als marginal betrachtet werden kann, also nicht zu einer Senkung des Marktpreises führt. Falls es zu einem Gewinnrückgang bei den konkurrierenden Anbietern käme, wäre die Outputsteigerung nicht in vollem Umfange ein Wohlfahrtsgewinn, sondern teilweise ein Wohlfahrtstransfer von den bisherigen Anbietern zum Job-Corps-Programm. Entsprechend wird die Arbeitslosigkeit von Teilnehmern zum Teil durch die Arbeitslosigkeit von Nichtteilnehmern ersetzt, falls es aufgrund der Produktion im Programm zu einer Angebotseinschränkung bei den konkurrierenden Anbietern kommt.

Die Verdrängungseffekte sind auf mikroökonomischer Ebene nicht korrekt erfaßbar, da das Marktergebnis, das sich ohne Programm eingestellt hätte, nicht bekannt ist. Sie spiegeln sich jedoch in der makroökonomisch orientierten Diskussion um die „natürliche Arbeitslosigkeit“ und die Phillipskurve wider. Einer der wenigen Versuche, die

aktive Arbeitsmarktpolitik mit dieser Diskussion zu verbinden, ist die Studie von Baily/Tobin (1978). Sie kommen zu dem Ergebnis, daß in einer unterbeschäftigten Wirtschaft zusätzliche Staatsausgaben für Arbeitsmarktpolitik nicht unbedingt zu einem Crowding-out privater Ausgaben führen müssen. Bei einer nachgiebigen Geldpolitik ist es ihrer Ansicht nach möglich, die Phillipskurve durch geeignete arbeitsmarktpolitische Maßnahmen zu verschieben. Dieses ist zum einen von der Auswahl der Teilnehmer und zum anderen davon abhängig, inwieweit die Entwicklung des allgemeinen Lohnniveaus durch die vermehrte Beschäftigung von Randgruppen beeinflusst wird. Ihre Studie gibt jedoch keine Auskunft über die Wirkungen einer qualifizierenden Arbeitsmarktpolitik.

Der Einfluß verschiedener Formen der aktiven Arbeitsmarktpolitik auf das Beschäftigungsniveau kann also bislang empirisch nicht genau erfaßt werden. Bei einem grundsätzlichen Konsens über die Notwendigkeit aktiver Arbeitsmarktpolitik sind jedoch ihre Erfolge bei der Umverteilung von Chancen, Einkommen und Arbeitsmarktrisiken ungleich wichtiger. Folglich sollten sich die Evaluationsanstrengungen auf die *strukturellen* Wirkungen der Arbeitsmarktpolitik konzentrieren, um die gewünschte Umverteilung möglichst effizient zu organisieren. Es gilt also Kriterien zu wählen, die hinreichend genau erfassen, welche Maßnahme bei gegebenen Kosten den höchsten Beitrag zum Ausgleich von Wettbewerbsnachteilen bietet.

Der überzeugendste Beweis für eine Integration am Arbeitsmarkt ist das Vorliegen eines Beschäftigungsverhältnisses. Unter Effizienzgesichtspunkten kann hierbei nicht die Beschäftigung innerhalb der Maßnahmen gezählt werden, sondern nur reguläre Beschäftigungsverhältnisse, die nach Abschluß der Maßnahme angetreten werden. Die in den Vereinigten Staaten häufig verwendeten Indikatoren „Arbeitszeit“ und „Anteil der Wochen mit Beschäftigung am Beobachtungszeitraum“ erscheinen aus europäischer Sicht nur wenig geeignet, da mit ihnen nicht erfaßt wird, ob es sich um ein dauerhaftes Arbeitsverhältnis oder eine Vielzahl von unterschiedlichen aufeinanderfolgenden Arbeitsverhältnissen handelt. Neben der Wahrscheinlichkeit des Eintritts in eine Beschäftigung sollte also auch die Art des Arbeitsverhältnisses (unbefristet, befristet, aushilfsweise) erfaßt werden. Die Beurteilung der Kategorie „befristet“ ist insofern schwierig, da gerade bei Angehörigen einer Problemgruppe Arbeitgeber die Möglichkeit einer Befristung des Arbeitsverhältnisses auf bis zu zwei Jahre als verlängerte Probezeit nutzen (vgl. Büchtemann/Höland 1989, S. 253 ff.). Wievielen Teilnehmern tatsächlich die Integration ins Arbeitsleben geglückt ist, kann folglich erst nach einer Beobachtungszeit von mehr als zwei Jahren nach Programmende festgestellt werden.

Eine längerfristige Beobachtung der Erwerbsbiographien der Maßnahmeteilnehmer ist auch aus anderen Gründen wünschenswert. Zum einen ist die Nachhaltigkeit der Maßnahmewirkungen ein entscheidendes Kriterium für ihren gesamtgesellschaftlichen Nutzen⁵. Zum anderen sollte überprüft werden, ob Kurzzeitmaßnahmen, die nur ausgewählte Qualifikationen vermitteln, einen längerfristigen Erfolg am Arbeitsmarkt sicherstellen können oder ob ihre Erfolge schneller als die von längerfristigen Maßnahmen erodieren. Für die Weiterentwicklung der Arbeitsmarktpolitik sind relativ kurzfristige Rückkoppelungen wichtig, doch sollte bekannt sein, ob sich der kurz- und längerfristige Erfolg verschiedener Maßnahmen parallel entwickelt oder nicht.

⁴ Zu ähnlichen Ergebnissen kommt auch Wiseman (1976).

⁵ In der Studie von Long/Mallar/Thornton (1981, S. 71 f.) zeigt die Sensitivitätsanalyse, daß der Gesamtnutzen vor allem von der Dauerhaftigkeit der Maßnahmewirkungen abhängt.

Ein anderer im Ausland häufig verwendeter Erfolgsmaßstab ist die Einkommensentwicklung. Für diesen Indikator sprechen zwei Überlegungen: zum einen steuert die erwartete Einkommensentwicklung nach den Vorstellungen der Humankapitaltheorie das Investitionskalkül der Individuen, und zum anderen ist sie ein Indikator für die Qualität des erlangten Arbeitsplatzes. Sofern keine Zufallsauswahl der Teilnehmer möglich ist, entsteht bei der Beurteilung der Einkommensentwicklung vor allem das Problem, wie die Bezugsgruppe zur Schätzung der hypothetischen Status-quo-Entwicklung zu wählen ist.

Eine Möglichkeit besteht darin, die Einkommensentwicklung der Programmteilnehmer mit der allgemeinen Einkommensentwicklung – unabhängig vom Einkommensniveau – zu vergleichen. Wie Ashenfelter (1978) zeigt, ist dabei die Wahl des Basiszeitraumes von entscheidender Bedeutung, da bei den Programmteilnehmern in der Regel unmittelbar vor Beginn der Maßnahme ein Einkommensrückgang zu verzeichnen ist. Dieser ist damit zu erklären, daß der Arbeitnehmer arbeitslos geworden ist und das Arbeitslosengeld unter dem vorherigen Einkommen liegt. Je nachdem, ob es sich nur um eine vorübergehende Minderung des Einkommens handelt, die nach einer Arbeitsaufnahme wieder behoben wäre (d. h. der Arbeitnehmer hätte ohne die Maßnahme einen gleich guten Arbeitsplatz gefunden) oder ob es sich um einen dauerhaften Rückgang handelt, der darauf zurückzuführen ist, daß die Qualifikationen des Arbeitnehmers am Markt weniger begehrt sind, ist der Bezugszeitpunkt für das Status-quo-Einkommen unterschiedlich zu wählen, um den Programmeffekt korrekt zu ermitteln.

Eine andere Möglichkeit besteht darin, eine Personengruppe zu betrachten, die zu einem bestimmten Zeitpunkt homogen war, etwa durch den Status „Arbeitslosigkeit“. Allgemeine Aussagen sind in diesem Fall nur möglich, wenn der Einfluß, den unterschiedliche Erwerbsbiographien vor Eintritt in die Arbeitslosigkeit auf die Einkommensentwicklung nach der Maßnahme haben können, berücksichtigt wird. Dies ist jedoch weniger von Bedeutung, wenn es um die Wirksamkeit verschiedener Maßnahmen in einer bestimmten Situation geht.

Als ein in der Aussage ähnlicher Indikator wird in den Vereinigten Staaten häufig die veränderte Abhängigkeit von Transferzahlungen einbezogen. Wenn auch nach Maßnahme noch Transferzahlungen notwendig sind, zeigt dies, daß die Einkommenssteigerung bei den Teilnehmern nicht zur Erlangung des allgemein anerkannten Mindestlebensstandards ausreicht. Dies kann als Ausdruck einer noch nicht ausreichend gestärkten Wettbewerbsfähigkeit interpretiert werden. Bei Verwendung dieses Indikators kommt es jedoch zu Verzerrungen, sofern in einer Maßnahme neue Ansprüche auf Arbeitslosengeld erworben werden können.

Fazit unserer Betrachtungen ist, daß sich gegenüber einer ausschließlichen Verwendung der Zahl der geschaffenen Arbeitsplätze als Erfolgsmaßstab der Arbeitsmarktpolitik zusätzliche Informationen durch die Einbeziehung der Entwicklung von Transferzahlungen und Einkommen gewin-

nen lassen. Die Einkommensentwicklung ist dabei tendenziell der aussagekräftigste Indikator. Dabei hängt die Höhe des Einkommenseffektes stark von der Spezifikation des zugrundeliegenden Modells ab, nicht aber der relative Erfolg verschiedener Maßnahmen im Vergleich (vgl. Barrow 1987). Um eine generell befürwortete Arbeitsmarktpolitik effizient zu gestalten, reicht es jedoch aus, für jede Arbeitsmarktsituation die geeignetste Maßnahme oder die bestmögliche Ausgestaltung identifizieren zu können. Um diesen Vergleich durchführen zu können, ist es unabdingbar, die verschiedenen Maßnahmen anhand desselben Erfolgsmaßstabes zu evaluieren.

Bei einem solchen Vergleich sollten auch die unterschiedlichen Kosten berücksichtigt werden, da eine geringere Erfolgswahrscheinlichkeit einer Maßnahme unter Umständen durch hinreichend niedrigere Programmkosten ausgeglichen wird. Ein Beispiel für eine solche Maßnahme sind Prämienzahlungen an Arbeitslose, die innerhalb kurzer Zeit eine neue Stelle finden (vgl. Woodbury/Spiegelman 1987). Bei diesem Experiment in Illinois wurde einer Gruppe von Arbeitslosengeldempfängern eine Bonuszahlung von 500 Dollar für den Fall versprochen, daß sie innerhalb von elf Wochen nach dem Verlust des alten Arbeitsplatzes einen neuen finden würde. Eine andere Gruppe diente als Kontrollgruppe und erhielt nur die üblichen Leistungen für Arbeitslose. Die Beschäftigungswahrscheinlichkeit der Experimentalgruppe stieg geringfügig an. Ihr Einkommen entwickelte sich nicht signifikant anders als das der Kontrollgruppe. Dies läßt den Schluß zu, daß der bessere Erfolg der Experimentalgruppe auf verstärkte Suchaktivitäten und nicht auf eine Senkung des verlangten Einstellungslohnes zurückzuführen ist⁶.

Die Kosten dieser Maßnahme betragen weniger als 500 Dollar pro Erfolgsfall, da nicht alle Berechtigten ihre Gutscheine einlösten. Für die erfolglosen Teilnehmer waren die Kosten der Arbeitsverwaltung Null. Andere Maßnahmen mit vermutlich geringen Kosten pro Teilnehmer sind Bewerbungstraining und Anleitung zur strukturierten Stellensuche⁷. Für derartige Maßnahmen mit geringen Kosten pro Teilnehmer können die Zugangskriterien recht großzügig festgelegt werden. Teurere Maßnahmen mit einem höheren Wirkungsgrad sollten für jene Arbeitslose verwendet werden, die größere Eingliederungsprobleme haben und für die daher die Maßnahmen mit geringer Eingriffintensität nicht ausreichen.

Das kanadische Job Development Program für Langzeitarbeitslose ist ein Beispiel dafür, wie Maßnahmen mit unterschiedlichen Kosten auf verschiedene Zielgruppen zugeschnitten werden können (vgl. GOSS, Gilroy & Associates 1989). In diesem Programm wurden Ausbildung und subventionierte Tätigkeiten miteinander verbunden. Dabei standen drei Optionen zur Verfügung, die unterschiedliche Anforderungen in bezug auf die Ausgestaltung der Tätigkeit und des Qualifikationsanteils vorsahen, aber auch mit unterschiedlich hohen Lohnsubventionen verbunden waren. Die Langzeitarbeitslosen wurden je nach der anhand der beobachtbaren Merkmale vermuteten Wettbewerbsschwäche in drei Gruppen eingeteilt. Für Langzeitarbeitslose ohne weitere Vermittlungshemmnisse betrug die Subvention sechzig Prozent. Für Behinderte, entlassene Strafgefangene und Angehörige ethnischer Minderheiten wurden zunächst achtzig, dann fünfzig und schließlich fünfundzwanzig Prozent Lohnsubvention gezahlt, die Unternehmen mußten einen individuellen Ausbildungsplan für die Teilnehmer entwerfen. In der dritten Alternative wur-

⁶ Ein anderes interessantes Ergebnis des Experimentes ist, daß Bonuszahlungen an Arbeitslose relativ erfolgreicher als gleichhohe einmalige Lohnzuschüsse an die Arbeitgeber sind (vgl. Woodbury/Spiegelman 1987, S. 520 ff.).

⁷ Ein Indiz dafür, daß diese Maßnahmen hohe Ertrags-Kosten-Relationen aufweisen, ist darin zu sehen, daß sie für Hochschulabgänger als private Dienstleistungen angeboten werden.

den Arbeitslose untergebracht, die nach Ansicht der Arbeitsbehörde aufgrund motivationaler Probleme nicht vermittelbar waren. Hier wurden nicht nur die vollen Lohnkosten erstattet, sondern darüber hinaus auch Zuschüsse für die Organisations- und Sachkosten geleistet. Hinter diesen hohen Zuschüssen stand die Überlegung, daß die Teilnehmer nur durch praktische Tätigkeiten langsam an den Arbeitsmarkt herangeführt werden könnten.

Alle Projektvarianten waren ausreichend hoch dotiert, so daß bei den Unternehmen keine Akzeptanzprobleme entstanden. Sie waren alle in dem Sinne erfolgreich, daß sie zu einer Steigerung der Beschäftigungswahrscheinlichkeit führten, die Lohnsteigerungen waren weniger beeindruckend. Es zeigte sich, daß die Programme in den Regionen, in denen zusätzlich theoretischer Unterricht von den Arbeitsämtern angeboten oder die Arbeitsplatzsuche durch Job-search-clubs unterstützt wurde, besonders erfolgreich waren. Dies gibt einen Hinweis darauf, daß Synergieeffekte zwischen Maßnahmen verschiedener Eingriffssintensitäten bestehen.

Folglich sollten Evaluationen der Arbeitsmarktpolitik in Zukunft vermehrt Maßnahmen miteinander vergleichen und ihre Wechselwirkungen untersuchen, um eine Feinsteuerung der Maßnahmen zu ermöglichen. Um auch Ergebnisse verschiedener Evaluationen in diesem Prozeß nutzen zu können, sollten die Zielfunktionen möglichst ähnlich spezifiziert werden. Dies ist insbesondere dann notwendig, wenn neben der Beschäftigungswahrscheinlichkeit die Einkommensentwicklung berücksichtigt wird, um weitere Informationen zu gewinnen. Allein mit dem Vergleich von verschiedenen Maßnahmen und der Untersuchung ihrer Wechselwirkungen kann jedoch noch keine Aussage über den optimalen Programm mix gemacht werden, da hierzu auch die optimale Größe der Programme bestimmt werden muß. Wie dieses Problem von der Evaluierung unterstützt werden kann, zeigt das folgende Kapitel.

3 Durchschnitts- versus Marginalanalyse in Evaluationen

Evaluationen, die auf eine Schätzung der optimalen Kapazität eines arbeitsmarktpolitischen Programms hinzielen, müssen zwei Fragen beantworten können: Zum einen sollen sie Unterschiede zwischen der tatsächlichen Entwicklung mit Programm und der hypothetischen Entwicklung ohne Programm („Status-quo-Entwicklung“) möglichst quantitativ schätzen. Zum anderen sollen sie feststellen, wie die Ertragsraten des Programmes durch eine Erweiterung oder Einschränkung beeinflußt werden.

Die beiden Probleme sollen an der folgenden Zielfunktion einer Qualifizierungsmaßnahme, die beispielhaft für das Vorgehen der Evaluationsforschung ist, verdeutlicht wer-

$$(1) \quad Y_i = \alpha + \beta \cdot T_i + \gamma \cdot X_i + U_i$$

Dabei bezeichnet Y_i das Einkommen, T_i eine Dummy-Variable, die für Programmteilnehmer den Wert 1 und ansonsten 0 annimmt, X_i einen Vektor der Charakteristika der Teilnehmer und U_i den Störterm, der sowohl die Meßfehler als auch die Zufallsfehler umfaßt.

⁸ Es gibt ökonometrische Verfahren zur Kontrolle der beobachtbaren und unbeobachtbaren Heterogenität zwischen den Teilnehmern und Nichtteilnehmern. „These techniques rely, however, on fundamentally untestable assumptions about the relationships among the unobservable variables, the selection process, and the outcomes, and are therefore of unknown efficacy in any particular application“ (Burtless/Orr 1986, S. 612).

Wenden wir uns zuerst den Lösungen für eine zuverlässige Schätzung der Status-quo-Entwicklung zu. Der vor allem wegen gesetzlicher und finanzieller Restriktionen oft verwendete nicht-experimentelle Ansatz versucht die Kontrollgruppe so zu bilden, daß ihre Mitglieder und die Programmteilnehmer im Hinblick auf die das Einkommen beeinflussenden Eigenschaften X_i wie beispielsweise Alter, Geschlecht, Schulbildung, Ausbildung, Beruf, Betriebszugehörigkeitsdauer, Branche und Region möglichst homogen sind. Auf diese Weise können für Teilnehmer mit differierenden Merkmalen unterschiedliche Ertragsraten für die Teilnahme an den Arbeitsmarktprogrammen ermittelt werden. So hat sich beispielsweise herausgestellt, daß die Erträge für Frauen in vielen US-Maßnahmen höher als für Männer sind. Wenn die X_i alle für die Einkommensentwicklung relevanten Einflußfaktoren umfassen, wären die Einkommensunterschiede ein korrektes Maß für den Erfolg der arbeitsmarktpolitischen Maßnahme.

In der Praxis läßt sich jedoch der theoretische Anspruch, X_i ohne Fehler zu messen, nicht einlösen. Tatsächlich können nur *beobachtbare* Merkmalsunterschiede erfaßt werden, dagegen lassen sich einige für den Einkommenszuwachs maßgebliche Charakteristika, wie beispielsweise die Motivation und das Selbstbewußtsein der Teilnehmer, nicht messen. Aufgrund dieser unbeobachtbaren Heterogenität, die zu einem im Umfang unbekanntem Erfolgsunterschied zwischen der Experimental- und der Kontrollgruppe führt, leidet der nicht-experimentelle Ansatz unter dem „selection bias“⁸. Zu vermuten ist dann, daß die Mitglieder der Experimentalgruppe aufgrund ihrer wahrscheinlich höheren Motivation auch ohne die Fördermaßnahme besser als die Kontrollgruppe auf dem Arbeitsmarkt abgeschnitten hätten, d. h. diese Methode überschätzt den Erfolg des Programms in einem unbekanntem Maße.

Im Design des in den USA favorisierten experimentellen Ansatzes kann ein „selection bias“ nicht entstehen, da die Bewerber nach einem Zufallsverfahren entweder der Experimentalgruppe, die die Förderung erhält, oder der Kontrollgruppe, der die Teilnahme verweigert wird, zugewiesen werden. Damit sind systematische Differenzen zwischen den beiden Gruppen aufgrund der unbeobachtbaren Charakteristika ausgeschlossen, so daß die Erfolgsunterschiede tatsächlich dem durchgeführten Programm zugerechnet werden können. Der Wert von Experimenten liegt darin, daß sie den Meßfehler aus U_i eliminieren. Die Kritik an ihnen basiert auf einem ethischen Argument: Ist es vertretbar, einem Teil der Bewerber die Teilnahme an dem Programm ausschließlich deswegen zu verweigern, damit eine wissenschaftliche Untersuchung durchgeführt werden kann? In den USA wird dieser Aspekt als ein letztlich unvermeidbares Opfer für die Entwicklung besserer zukünftiger arbeitsmarktpolitischer Programme angesehen. In Europa werden aus diesem Grunde nahezu keine Experimente durchgeführt.

Angesichts der heftigen Kontroversen über eine zuverlässige Methode zur Bildung der Kontrollgruppe ist ein – aus ökonomischer Perspektive ungleich bedeutenderer Aspekt übersehen worden, nämlich die methodische Herangehensweise zur Maximierung des Erfolgs eines Programms. In Gleichung (1) gibt β den *durchschnittlichen* Einkommenszuwachs bei den Teilnehmern aufgrund der Fördermaßnahme an. Als Politikempfehlung folgt aus diesem Ansatz: eine Maßnahme gilt als erfolgreich, sofern α hoch genug ist, um die Kosten des Programms pro Teilnehmer zu decken, andernfalls sollte sie abgebrochen werden

(z. B. Burtless/Orr 1986). Zur Maximierung des Erfolgs dürfen jedoch nicht die Durchschnittserträge und Durchschnittskosten, sondern es müssen die *Grenzerträge* und *Grenzkosten* miteinander verglichen werden. Da bei heterogenen Teilnehmern Durchschnitts- und Grenzkosten nicht übereinstimmen, führt die übliche ökonomische Vorgehensweise zu falschen Politikempfehlungen.

Zur Verbindung von ökonomischer Analysetechnik und ökonomischen Optimalbedingungen verwenden Björklund/Moffitt (1987) ein Random Coefficient Modell, in dem die Heterogenität der Teilnehmer durch individuelle Zufallsabweichungen von den durchschnittlichen Erträgen und Kosten eines Programms berücksichtigt werden. Sie unterscheiden dabei zwei Quellen von Heterogenität, die zu unterschiedlichen Entscheidungen über die Teilnahme an einem Programm führen: die Heterogenität in den Erträgen besagt, daß die Teilnehmer eines Qualifizierungsprojektes, beispielsweise eines Kurses über Textverarbeitung, unterschiedliche Einkommenssteigerungen aufgrund der Maßnahme erwarten; die Opportunitätskosten der Nichtteilnahme sind unterschiedlich hoch. Heterogenität in den Kosten oder Präferenzen besagt, daß die verschiedenen Teilnehmer unterschiedlich hohe Anstrengungen monetärer und nicht-monetärer Art auf sich nehmen müssen, um das gleiche arbeitsmarktpolitische Ziel, beispielsweise die gleiche Einkommenssteigerung, zu erreichen.

Wenn heterogene Präferenzen berücksichtigt werden, läßt sich feststellen, ob die Effizienz eines arbeitsmarktpolitischen Förderprogramms durch eine Veränderung seines Umfangs gesteigert werden kann. In der vorherrschenden ökonomischen Spezifikation wird die Programmwirkung durch den für alle Teilnehmer gleichen Koeffizienten a angegeben. Wenn das Programm bei einem einheitlichen a für einen Teilnehmer ineffizient ist, gilt diese Beurteilung automatisch für alle Teilnehmer, d. h. die wirtschaftspolitische Empfehlung kann nur Ja oder Nein lauten. Durch die Berücksichtigung heterogener Kosten und Erträge werden marginale Entscheidungen möglich. Da mit steigender Größe eines Programms Bewerber mit immer geringeren Ertragsersparungen hinzukommen, läßt sich über eine Steuerung des Zugangs die durchschnittliche Ertragsrate nach oben oder unten verändern. Auf diese Weise läßt sich der Umfang des Programms so steuern, daß die Nutzen-Kosten-Relation maximal oder zumindest positiv wird. Neben der Ja-Nein-Entscheidung wird damit die Bandbreite wirtschaftspolitischer Empfehlungen um eine Vielzahl gradueller quantitativer Kapazitätsanpassungen bereichert.

Zur korrekten Durchführung der Evaluation sollte zwischen nachfrage- und angebotsbeschränkten Programmen unterschieden werden (vgl. Björklund 1988). Bei den *angebotsrestringierten* Programmen gibt es mehr Bewerber als verfügbare Plätze. Die Arbeitsmarktpolitiker können in diesem Fall zwischen zwei Strategien wählen: entweder wird die Kapazität ausgeweitet, um die gesamte Nachfrage befriedigen zu können, oder es werden durch eine Senkung der Zuschüsse die Teilnahmekosten gesteigert, um die Nachfrage an die gegebene Kapazität anzupassen.

Damit keine Verzerrungen zwischen der Experimental- und der Kontrollgruppe auftreten, muß die zweistufige Auswahlprozedur für beide Gruppen gleich sein. Zunächst muß das Nutzenkalkül der Interessenten positiv ausgehen, d. h. ihre individuellen Ertragsersparungen müssen die individuellen Kostenerwartungen übersteigen, damit sie sich überhaupt bewerben. Sodann muß die Behörde nach

gesetzlichen Vorgaben oder nach Leistungskriterien über die Zulassung zu dem Programm entscheiden. Erst nach der Auswahl der Behörden erfolgt die durch ein Zufallsverfahren gesteuerte Verteilung der Teilnehmer auf Experimental- und Kontrollgruppe. Bei diesem Design muß der Kontrollgruppe die Teilnahme an dem Förderprogramm verweigert werden, obwohl sie die Zulassungsbedingungen der Behörde erfüllt.

Um zu vermeiden, daß trotz des Bewerberüberhangs die Kapazität der Fördermaßnahme teilweise ungenutzt bleibt, kann die Behörde mehr Teilnehmer als üblich auswählen, so daß der tatsächliche Umfang des durchgeführten Programms unverändert bleibt. Insofern haben Evaluationen von angebotsbeschränkten Programmen den Vorteil, daß das ethische Auswahlproblem nicht entsteht. Da nicht genug Plätze für alle Bewerber zur Verfügung stehen, muß auf jeden Fall eine Auswahl erfolgen. Unter diesen Umständen wird eine Zufallsauswahl häufig als ein gerechtes Zuteilungsverfahren angesehen.

Allerdings kommen bei einer so gestalteten Evaluation auch Teilnehmer zum Zuge, die normalerweise nicht die behördliche Zulassung bekommen hätten. Dies kann zur Heranführung der Evaluationstechnik an die mikroökonomischen Optimalbedingungen genutzt werden, wenn die Behörde die relativen Ertragsersparungen der potentiellen Teilnehmer ungefähr einschätzen kann. Die diesbezüglichen Anforderungen sind gering. Falls 250 Programmplätze vorhanden sind, genügt es, wenn die Behörde in einem ersten Schritt die 250 Personen mit den besten Erfolgsaussichten in eine erste Gruppe und die nächsten 250 Personen in eine zweite Gruppe einteilt. Eine genaue individuelle Reihenfolge ist nicht erforderlich. Notwendig ist lediglich, daß die im zweiten Schritt erfolgende Zufalls-einteilung in Experimental- und Kontrollgruppe für beide Personenkreise getrennt erfolgt. Das Ergebnis der besseren Teilnehmer spiegelt die *durchschnittlichen* Erträge des bisherigen Programms, das Ergebnis der zweiten Gruppe gibt die *marginalen* Erträge eines im Umfang verdoppelten Programms an (vgl. Björklund 1989).

Bei *nachfragerestringierten* Programmen gibt es eine hinreichende Zahl von Ausbildungsplätzen, so daß allein das Nutzenkalkül der Bewerber über eine Teilnahme entscheidet. Die Arbeitsmarktpolitiker können auch in diesem Fall zwischen zwei Strategien wählen: entweder werden die Teilnahmekosten über eine Erhöhung der Zuschüsse gesenkt, um die gegebene Kapazität auszuschöpfen, oder die Kapazität wird reduziert, um sie an die gegebene Nachfrage anzupassen.

Grundlage für eine Evaluation ist die Bildung einer Experimentalgruppe und einer Kontrollgruppe aus der Gesamtheit der Bewerber zu den anfänglich geltenden Bedingungen. Die Differenz zwischen den Erfolgen der beiden Gruppen zeigt die *durchschnittlichen* Erträge des Programms in seinem bisherigen Umfang an. Wird nun die Attraktivität des Programms durch eine Lockerung der Zugangsbedingungen erhöht, verändert sich seine ökonomische Vorteilhaftigkeit. Teilnehmer mit geringeren Ertragsersparungen kommen als zusätzliche Gruppe hinzu. Teilt man diese zusätzlichen Bewerber ebenfalls wieder in eine Experimental- und eine Kontrollgruppe auf, erhält man die *marginalen* Erträge des erweiterten Programms. Hieraus können dann ökonomisch korrekte Schlüsse über eine Erweiterung oder Einschränkung des Programms gezogen werden.

Wie wichtig die Strategie der quantitativen Anpassung für die Steigerung der Effizienz der aktiven Arbeitsmarktpolitik ist, zeigen die Ergebnisse von Björklund/Moffitt (1987) und Björklund (1988, S. 272). Bei der Evaluation eines schwedischen Qualifizierungsprogramms haben sie eine durchschnittliche Rendite von 11% und sogar von 18% für junge Arbeiter ohne Lehre geschätzt. Trotz dieses vergleichsweise großen Erfolgs war der *marginale* Ertrag für alle Teilnehmergruppen negativ. Dies zeigt, daß der Umfang des Programms überdimensioniert gewesen ist, d. h. die aktive Arbeitsmarktpolitik könnte in einem besseren Licht dastehen, wenn die Grenznachfrager nicht zugelassen worden wären. Im konkreten Fall schlagen die Autoren vor, die Unterhaltszahlungen an die Teilnehmer zu kürzen, um die potentiellen Bewerber mit den geringsten Ertragserwartungen per Selbstauswahl aus dem Programm herauszuhalten.

Evaluationen auf Basis von Marginalanalysen sind insbesondere für die europäische Arbeitsmarktpolitik relevant, denn sie zielen darauf ab, die vorhandenen Programmtypen durch marginale Änderungen optimal an die jeweilige Arbeitsmarktlage anzupassen. Dieser Vorgehensweise liegt der Konsens zugrunde, daß alle Formen der aktiven Arbeitsmarktpolitik für irgendeine Problemgruppe zu irgendeiner Zeit nützlich sein können und sind. Deshalb hat in Europa das Ziel Priorität, die richtige Zielgruppe für die verschiedenen Programme zu finden. Die Bundesanstalt für Arbeit könnte viel Kritik an ihren Maßnahmen vermeiden, wenn sie bei dem Umfang ihrer Programme und der Auswahl der Teilnehmer selektiver vorgehen würde. Je besser es den Arbeitsmarktbehörden gelingt, die für die jeweilige Maßnahme geeigneten Bewerber auszuwählen, um so positiver wird die Evaluation der Maßnahme ausfallen. Das hier geschilderte und an einem schwedischen Qualifizierungsprogramm erprobte Evalu-

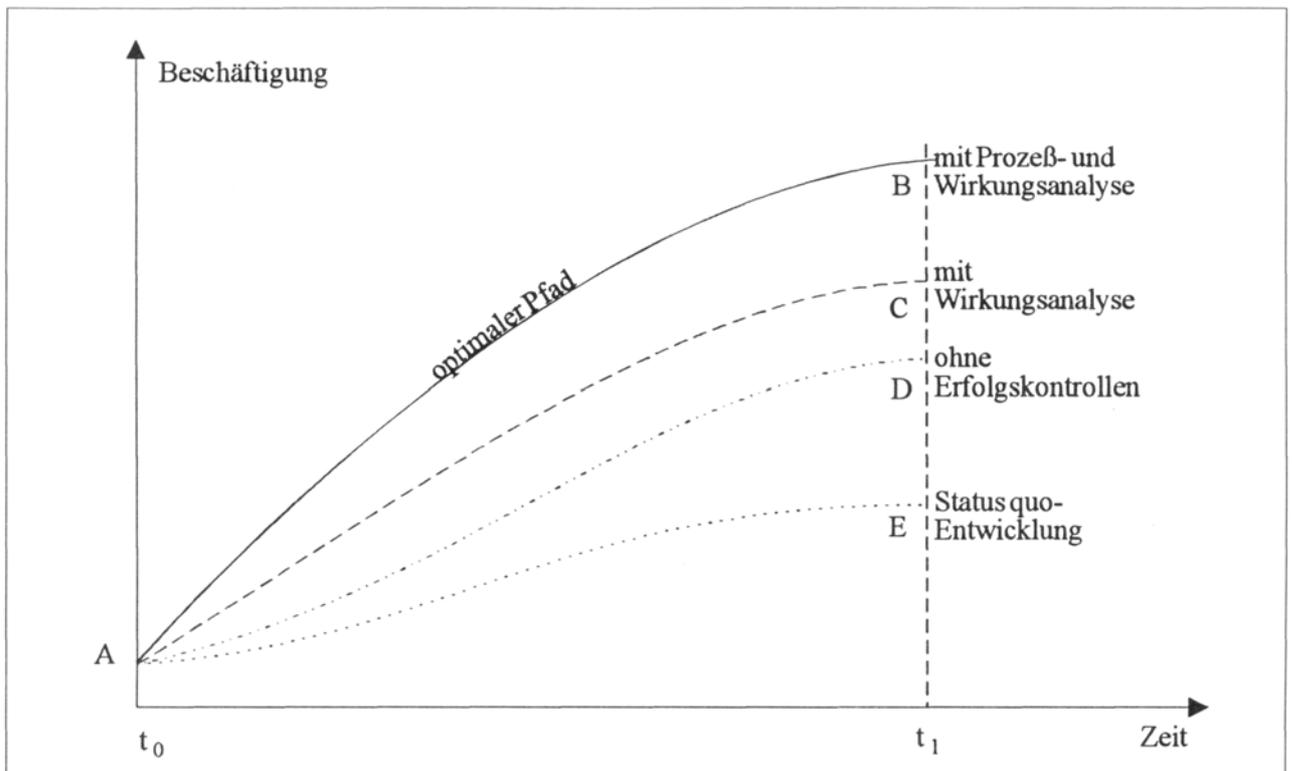
ierungsverfahren erlaubt es, eine der wesentlichen Ursachen für ineffiziente Programme zu beseitigen.

4 Von der Wirkungsanalyse zur Prozeßanalyse

Der Erfolg einer arbeitsmarktpolitischen Maßnahme ist nicht nur an Verbesserungen gegenüber der Situation ohne Maßnahme zu messen, sondern es sollte auch untersucht werden, ob alle zur Verfügung stehenden Möglichkeiten ausgeschöpft worden sind. Wenn bislang Chancen ungenutzt geblieben sind, kann dies sowohl an einer unzureichenden Umsetzung des akkumulierten Erfahrungswissens als auch an fehlenden Möglichkeiten zur Nutzung aktuellen Wissens liegen. Bei einer vollen Nutzung aller Chancen würde man den optimalen Entwicklungspfad *AB* in Abb. 1 erreichen. Dazu bedarf es eines koordinierten Zusammenwirkens von Wirkungs- und Prozeßanalyse.

Die Prozeßanalyse ist eine relativ neue Evaluationsmethode, so daß es noch unterschiedliche Vorstellungen über ihre Aufgaben gibt. Oft wird der wesentliche Unterschied zwischen den beiden Analysemethoden darin gesehen, daß eine Wirkungsanalyse die Wirkungen eines gegebenen Programms oder eines gegebenen Instruments in einer quantitativen Weise angibt, während eine Prozeßanalyse den institutionellen Rahmen eines Programms in einer qualitativen Weise überprüft. Demgemäß weist Hasan (1991, S. 10) als Resümee einer Expertenkonferenz der OECD der Prozeßanalyse zwei Funktionen zu: sie soll zum einen in der Form des „monitoring“ prüfen, ob das Programmdesign die politischen Ziele der Arbeitsmarktpolitik widerspiegelt. Hierzu werden die Output- und Inputgrößen eines Programms erfaßt, beispielsweise die Zahl der Teilnehmer, die angebotenen Förderungstypen und ihre -intensitäten sowie möglicherweise ein Vergleich mit vorgegebenen Zielgrößen durchgeführt. Zum anderen soll sie in der Form

Abbildung 1: Nutzung vorhersehbarer und nicht-vorhersehbarer Informationen



der Implementationsstudie prüfen, ob der verwaltungsmäßige Ablauf eines Programms optimiert ist. Sie soll u. a. herausfinden, wie die Behörden in einem dezentralisierten Gemeinwesen das Programmziel interpretieren und ob durch einen Vergleich der Implementation in verschiedenen Verwaltungsbezirken örtliche Verbesserungen im Verwaltungsablauf generalisierbar sind.

Ein anderes Abgrenzungskriterium wählt Nathan (1991). Er sieht den Schwerpunkt der Wirkungsanalyse bei den individuellen und den Schwerpunkt der Prozeßanalyse bei den institutionellen Wirkungen. Die institutionellen Aspekte umfassen zum einen die Anreizwirkungen innerhalb der Administration. Schwierigkeiten entstehen in den USA – aber keineswegs nur dort – vor allem dadurch, daß die gesetzgeberische Gestaltungsfreiheit, die finanzielle Zuständigkeit und die Durchführung der Maßnahmen der aktiven Arbeitsmarktpolitik auf die drei Ebenen Bundesstaat, Staat und „county“ aufgeteilt sind. Aus politischer Sicht ist eine Ausgestaltung der institutionellen Regelungen wichtig, die dazu führt, daß die Behörden das Arbeitsmarktprogramm im Sinne der Gesetzesinitiatoren ausführen („principal-agent-problem“).

Zum anderen umfassen die institutionellen Aspekte die Anreizwirkungen bei den Adressaten der Maßnahme. Die beiden wichtigsten Steuerungsinstrumente gegenüber den Arbeitslosen oder von Arbeitslosigkeit bedrohten Individuen sind die Zulassungsbedingungen zu den Programmen und die finanziellen Unterstützungen. Eine gute institutionelle Ausgestaltung zeichnet sich dadurch aus, daß jeder Bewerber durch Selbstauswahl das für ihn maßgeschneiderte Programm präferiert. Adressaten der Arbeitsmarktpolitik sind aber auch die Unternehmen, die beispielsweise Lohnzuschüsse erhalten. Hier gilt es, institutionelle Regelungen zu finden, die verhindern, daß jeweils nur die wettbewerbsstärksten Mitglieder einer Zielgruppe eingestellt werden.

Beide Versuche einer Funktionsbeschreibung der Prozeßanalyse können nicht überzeugen. Hasans (1991, S. 10) Unterscheidung in eine mehr qualitative bzw. mehr quantitative Forschungsmethode wird von ihm schon dadurch eingeschränkt, daß er das weitgehend quantitative „monitoring“ ebenfalls der Prozeßanalyse zurechnet. Nathans (1991, S. 34) Unterscheidung in individuelle und institutionelle Wirkungen läßt ebenfalls eine scharfe Trennlinie zwischen den Analysemethoden vermissen. Theoretisch könnte man in der Wirkungsanalyse jede institutionelle Änderung als eine eigene Projektvariante interpretieren, entsprechende Experimental- und Kontrollgruppen bilden und die Wirkungsunterschiede aufgrund der institutionellen Variation schätzen. Wenn auch in der Praxis solche Wirkungsanalysen wegen des erforderlichen Zeit- und Ressourcenaufwandes nicht durchgeführt werden, gilt doch, daß auch institutionelle Änderungen letzten Endes stets durch individuelle Nutzensteigerungen gerechtfertigt werden müssen.

Als Fazit läßt sich festhalten, daß es die Intention sowohl der Wirkungs- als auch der Prozeßanalyse ist, Erfahrungen aus vergangenen Projekten für ein effizienteres Design zukünftiger Programme zu nutzen. Beide Evaluationsformen verfolgen das gleiche Ziel, nämlich die Effizienz von Arbeitsmarktprogrammen zu verbessern. Angesichts dieses Tatbestandes ist es wenig hilfreich, der Prozeßanalyse die qualitativen bzw. institutionellen Aspekte und der Wirkungsanalyse die quantitativen bzw. individuellen Aspekte zuzuweisen. Der entscheidende Unterschied zwischen den

beiden Evaluationsmethoden liegt unseres Erachtens auf einem bisher gänzlich vernachlässigten Gebiet.

Die unbestreitbaren Erfolge der Wirkungsanalysen haben die Arbeitsmarktpolitiker für die Tatsache blind gemacht, daß ökonomische Handlungskoordination in erster Linie ein *Informationsproblem* ist. Entscheidungen über Arbeitsmarktpolitik müssen stets unter Unsicherheit getroffen werden, d. h. selbst mit vollkommenen Wirkungsanalysen können nicht alle bei der Durchführung eines Programms auftretenden Probleme und Chancen vorausgesehen werden. Es ist das Problem der Verwertung von Wissen, das niemandem in seiner Gesamtheit gegeben ist. Folglich gibt es zum einen eine Teilmenge des gesamten in der Volkswirtschaft vorhandenen Wissens, das zentralisierbar und daher allgemein verfügbar ist. Zum anderen gibt es das Wissen über aktuelle Geschehnisse, das wegen seiner Kurzlebigkeit nicht zentralisierbar und deswegen auch nicht allgemein verfügbar ist. Um den optimalen Entwicklungspfad in Abb. 1 zu realisieren, müssen jedoch beide Teilmengen des Wissens in der Arbeitsmarktpolitik genutzt werden. Hierbei übernehmen Wirkungs- und Prozeßanalyse unterschiedliche Funktionen.

Die der Wirkungsanalyse zugrundeliegende neoklassische Theorie unterstellt, daß das für die Koordination der ökonomischen Aktivitäten erforderliche Wissen bekannt oder zumindest leicht erlangbar ist. Folglich soll es Aufgabe der Wirkungsanalyse sein, das für die Steuerung des Angebots an und der Nachfrage nach Arbeitsmarktmaßnahmen im Entscheidungszeitpunkt noch nicht bekannte, aber erforderliche Wissen zu bringen. Hierbei geht es um die Auswahl unter konkurrierenden Arbeitsmarkttheorien, unter konkurrierenden Instrumenten und unter konkurrierenden Annahmen über die exogenen Entwicklungen. Die Empirie soll also die Zweifelsfragen klären, die durch logische Deduktion allein nicht gelöst werden können.

In Abb. 1 soll der Entwicklungspfad *AD* den Beschäftigungserfolg eines Programms, das ohne Nutzung früherer Wirkungsanalysen geplant wird, angeben. Erfahrungsgemäß bleiben dabei Chancen ungenutzt, d. h. der optimale Entwicklungspfad *AB* wird nicht erreicht. Mit jeder Wirkungsanalyse werden einige bisherige Irrtümer über Kausalitäten revidiert und Wissenslücken über exogene Daten geschlossen. Im Idealfall kennt der Entscheidungsträger nach einer hinreichend großen Anzahl von Wirkungsanalysen die Präferenzen, Reaktionen und Budgetrestriktionen der Anbieter und Nachfrager auf dem Arbeitsmarkt so gut, daß er ein optimales Arbeitsmarktprogramm *konzipieren* kann.

Die weitere Entwicklung wird sich jedoch nur dann entlang des optimalen Pfades *AB* vollziehen, falls es der planenden Behörde gelungen ist, sämtliche während der Durchführungsphase auftretenden Ereignisse hinsichtlich Art, Umfang und Zeitpunkt exakt vorauszusehen und einzuplanen. In der Wettbewerbstheorie wird jedoch vehement bestritten, daß es jemals möglich wäre, alle Determinanten der Angebots- und Nachfragekurven *ex ante* empirisch zu erfassen. Das unternehmerische Tun beschränkt sich nicht auf das rein mechanische Berechnen des Schnittpunktes zwischen der Angebots- und der Nachfragekurve, da Verlauf und Lage dieser Kurven erst durch die Findigkeit der Unternehmer entdeckt werden müssen. Dementsprechend kann auch in der Arbeitsmarktpolitik der maximale Programmserfolg nur realisiert werden, wenn neben dem Erfahrungswissen aus vergangenen Projekten das aktuelle Wissen des laufenden Programms verwertet wird.

Das folgende Beispiel illustriert die Notwendigkeit, auf unvorhergesehene Ereignisse zu reagieren (vgl. Nathan 1991, S. 36 ff.). Im Jahre 1985 setzte Kalifornien das ehrgeizige Programm „Greater Avenues for Independence“ („GAIN“) in Kraft, das Sozialhilfeempfängern zu einem Arbeitsplatz verhelfen sollte. Das Programm sah eine bestimmte Abfolge von Eignungstests, Schulbildung, beruflicher Erst- oder Weiterbildung und Hilfe bei der Arbeitsplatzsuche vor und sollte von den „counties“ organisiert werden. Zwei unerwartete Ereignisse erforderten eine erhebliche Modifikation der ursprünglichen Pläne: zum einen hatte man mit einer Analphabetenquote von 25 Prozent gerechnet, tatsächlich gehörten dieser Gruppe jedoch 50 Prozent der Teilnehmer an. Diese Teilnehmerstruktur erforderte eine völlige Neudimensionierung der einzelnen Programmkomponenten, was nur durch eine hohe Flexibilität bei den rechtlichen Vorschriften, bei der Verwendung der sachlichen Mittel sowie bei dem Einsatz der Mitarbeiter möglich war. Zum anderen wurden die Budgets für „GAIN“ nach einem Jahr beträchtlich gekürzt. Auch hier zeigten sich starke Unterschiede in der Umsetzung der Budgetrestriktion: einige „counties“ stellten Sozialarbeiter ein, die auf einer kontinuierlichen Basis die Teilnehmer betreuten. Andere „counties“ sparten bei dieser Komponente, indem sie die Betreuung auf fachlichen Rat in speziellen Situationen beschränkten.

Dieses Beispiel zeigt, daß es bei jedem Programm Überraschungen gibt, die ad hoc Anpassungen verlangen. So geläufig jedem Beteiligten das Faktum der Überraschungen ist, so wenig hat es bisher Eingang in die Evaluationsforschung gefunden. Nach wie vor vermittelt das Studium der Literatur den Eindruck, daß man durch eine größere Anzahl von Wirkungsanalysen die Überraschungen vermindern oder gar ausschließen könne. Diese Vorstellung hat von Hayek (1976) als Anmaßung von Wissen abgelehnt. Was wir tatsächlich durch gute Wirkungsanalysen erreichen können, ist, daß wir statt des Beschäftigungserfolges AD die höhere Beschäftigungssteigerung AC realisieren können. Der Unterschied zwischen AD und AC besagt, daß im Vergleich zu einer Programmplanung ohne jegliche Erfolgskontrolle bereits eine Verbesserung aufgrund des Erfahrungswissens aus Wirkungsanalysen im Umfange von CD erreichbar ist. Aufgrund von Wirkungsanalysen allein ist es jedoch ausgeschlossen, bereits den optimalen Entwicklungspfad AB zu realisieren. Dazu bedarf es der zusätzlichen Nutzung aktuellen Wissens.

Aktuelles Wissen erlangen die Arbeitsmarktbehörden aus den unvorhergesehenen Überraschungen. Im GAIN-Programm haben sie aus dem großen Zustrom von Analphabeten gelernt, daß die Nachfragekurve nach schulischer Ausbildung viel weiter rechts als antizipiert gelegen hat. Insofern war dies der Beweis, daß am Anfang die Koordination zwischen den beiden Arbeitsmarktseiten noch höchst unvollkommen gewesen ist. Notwendigkeit zur Anpassung der Pläne besteht aufgrund der Entdeckung der Marktteilnehmer, daß ihre Erwartungen bezüglich der Erwartungen anderer zu optimistisch oder zu pessimistisch gewesen sind. Nur wenn dieses aktuelle Wissen genutzt wird, kann man auf den optimalen Entwicklungspfad AB gelangen.

Notwendige Voraussetzung für die Nutzung dieses an einem unbekanntem Ort, zu einem unbekanntem Zeitpunkt

⁹ So ist zum Beispiel Scharpf (1982, S. 308) der Ansicht, daß die Wirkungen des Sonderprogrammes für Regionen mit besonderen Beschäftigungsproblemen durch die unterschiedliche Implementation in den einzelnen Arbeitsamtsbezirken gesteigert werden konnte.

und einem unbekanntem Individuum zufallenden aktuellen Wissens ist Flexibilität bei der Durchführung der Arbeitsmarktmaßnahme. Wegen seiner Kurzlebigkeit ist aktuelles Wissen nicht zentralisierbar. Aus diesem Grunde kann es nicht als Ergebnis von Wirkungsanalysen in den generellen institutionellen Rahmen der Arbeitsmarktpolitik eingehen, sondern es kann erst vor Ort genutzt werden. Es ist Aufgabe der Prozeßanalyse, zu prüfen, ob dieses sich erst während der Durchführung der Arbeitsmarktmaßnahme offenbarende aktuelle Wissen, das zudem noch unter viele Personen aufgeteilt ist, genutzt wird. Ein Zeichen für die Verwertung aktuellen Wissens dürfte die regional unterschiedliche Implementation eines bundesweiten Programms sein, da sich die Arbeitsmarktlage, die Struktur der Arbeitslosen, das Arbeitsplatzangebot, die angebotenen und gesuchten Qualifikationen regional unterscheiden. Die Prozeßanalyse soll untersuchen, ob der institutionelle Rahmen der Arbeitsmarktpolitik den örtlichen Arbeitsämtern hinreichend große Flexibilität läßt, damit sie ihr Wissen über die regionalen Besonderheiten in die Programmdurchführung einbringen können⁹.

Zusammenfassend läßt sich festhalten, daß für eine effiziente Steuerung der Arbeitsmarktpolitik sowohl die Wirkungsanalyse als auch die Prozeßanalyse notwendig sind. Wirkungsanalysen nutzen das Erfahrungswissen aus erfolgreichen und erfolglosen vergangenen Projekten, um mit diesem zentralisierten Wissen das neue Programm zu planen. Insofern ist der Nutzen von *Wirkungsanalysen* darin zu sehen, die für das neue Programm *vorhersehbaren Wissensenselemente*, d. h. die zeitlich stabilen Rahmendaten zu liefern. Dagegen ist es Aufgabe der Prozeßanalyse, die *Verwertung des unvorhersehbaren aktuellen Wissens* zu ermöglichen. Insofern untersucht die Prozeßanalyse in erster Linie, wie die mikroökonomischen Kenntnisse über Anreizmechanismen Eingang in das Programmdesign finden können, damit die Beteiligten ihre Freiheitsräume im Sinne der gesamtwirtschaftlichen Effizienz nutzen. Mit dieser Aufgabenbeschreibung sind der Wirkungs- und der Prozeßanalyse klar voneinander unterscheidbare Funktionen zugewiesen worden.

5 Schluß

Das Studium der internationalen Evaluationsforschung zeigt, daß die USA führend in der theoretischen Entwicklung und praktischen Anwendung experimenteller Wirkungsanalysen sind. Trotz der forschungspolitischen Führungsrolle ist der arbeitsmarktpolitische Erfolg der amerikanischen Wirkungsanalysen eher bescheiden. Wilensky (1985) vermutet, daß dies auf eine Wechselwirkung von Politik und Evaluationen zurückzuführen ist. Da die amerikanische Arbeitsmarktpolitik auf relativ kurzfristigen Programmen beruht, sind Wirkungsanalysen notwendig, die eine klare Entscheidung für oder gegen ein Programm möglich machen. Unter den kritischen Augen der Öffentlichkeit kommt es zu einer Verbesserung der Evaluationsmethoden, die darauf zielen, den Fehler ersten Grades (das Programm wird fortgeführt, obwohl es ineffizient ist) zu vermeiden. Angesichts der in der Regel vergleichsweise geringen Einkommenszuwächse bei den Programmteilnehmern werden die verbleibenden methodischen Probleme und Beurteilungsspielräume der Evaluationen von den Gegnern aktiver Arbeitsmarktpolitik genutzt, um die Programme zu kürzen oder zu beenden.

Hingegen ist in Deutschland ein breiter politischer Konsens über die Notwendigkeit aktiver Arbeitsmarktpolitik festzu-

stellen. In diesem Kontext nehmen Wirkungsanalysen, die ein Programm rechtfertigen oder ablehnen sollen, nicht die höchste Dringlichkeitsstufe ein. Die positive Grundstimmung für eine aktive Arbeitsmarktpolitik kann jedoch nicht darüber hinwegtäuschen, daß Art und Umfang der Förderung der am Arbeitsmarkt Benachteiligten außerordentlich kontrovers diskutiert wird. An diesem Punkt muß die wissenschaftliche Beratung der Arbeitsmarktpolitik in Deutschland ansetzen. Hierzu sollten vergleichende Wirkungsanalysen verschiedener Maßnahmen unter Berücksichtigung der methodischen Erkenntnisse aus den Vereinigten Staaten durchgeführt werden, um den *relativen* Programmserfolg zu bestimmen. Die optimale *Größe* der Programme kann bestimmt werden, wenn die schwedischen Ansätze zur Berechnung marginaler Ertragsraten aufgenommen werden.

Die wichtigste Aufgabe sehen wir in der Weiterentwicklung der Prozeßanalyse. Die vorwiegend qualitative Prozeßanalyse untersucht, ob die institutionellen Regelungen der Arbeitsmarktpolitik so flexibel sind, daß nicht nur das Erfahrungswissen aus vergangenen Projekten, sondern auch das aktuelle Wissen des laufenden Programms zur Steigerung der Effizienz verwertet werden kann. Mit ihrer Hilfe läßt sich die Feinsteuerung der Instrumente hinsichtlich Adressatenkreis, Leistungsumfang und Differenzierung der Zuschüsse als Reaktion auf unvorhergesehene Überraschungen während der Programmdurchführung optimieren. Auch im Ausland wird die Notwendigkeit gesehen, diese Forschungsrichtung zu stärken. Dazu bieten die deutschen Erfahrungen mit der Implementationsanalyse einen guten Ansatzpunkt.

Literaturverzeichnis

- Ashenfelter, O. (1978): Estimating the Effect of Training Programs on Earnings. In: *The Review of Economics and Statistics*, Vol. 60, S. 47-57
- Baily, M. N./Tobin, J. (1978): Inflation-Unemployment Consequences of Job Creation Policies. In: J. L. Palmer (Hrsg.), *Creating Jobs: Public Employment Programs and Wage Subsidies*. Washington DC, S. 43-76
- Barnow, B.S. (1987): The Impact of CETA Programs on Earnings – A Review of the Literature. In: *The Journal of Human Resources*, Vol. 22, S. 157-193
- Björklund, A. (1988): What Experiments are needed for Manpower Policy? In: *The Journal of Human Resources*, Vol. 23, S. 267-277
- Björklund, A. (1989): Evaluations of Training Programs: Experiences and Proposals for Future Research. Discussion paper no. FS I 89-13 des Wissenschaftszentrums Berlin
- Björklund, A. (1991): Evaluation of Labor Market Policy in Sweden. In: OECD (Hrsg.), *Evaluating Labour Market and Social Programmes – The State of a Complex Art*. Paris, S. 73-88
- Björklund, A./Moffitt, R. (1987): The Estimation of Wage Gains and Welfare Gains in Self-Selection Models. In: *The Review of Economics and Statistics*, Vol. 69, S. 42-49
- Bloom, H. S. (1987): Lessons from the Delaware Dislocated Worker Pilot Program. In: *Evaluation Review*, Vol. 11, S. 157-177
- Bruche, G./Reissert, B. (1985): Die Finanzierung der Arbeitsmarktpolitik – System, Effektivität, Reformansätze. Frankfurt (usw.): Campus
- Bryant, E. C./Rupp, K. (197): Evaluating the Impact of CETA on Participant Earnings. In: *Evaluation Review*, Vol. 11, S. 473-492
- Burtless, G. (1985): Are Targeted Wage Subsidies Harmful? Evidence from a Wage Voucher Experiment. In: *Industrial and Labor Relations Review*, Vol. 39, S. 105-114
- Burtless, G./Orr, L. L. (1986): Are Classical Experiments needed for Manpower Policy? In: *The Journal of Human Resources*, Vol. 21, S. 606-639
- Büchtemann C. F. (unter Mitarbeit von A. Höland) (1989): Befristete Arbeitsverträge nach dem Beschäftigungsförderungsgesetz 1985 (BeschFG 1985). Band 183 der Reihe „Forschungsberichte“, hrsg. vom Bundesminister für Arbeit und Sozialordnung, Bonn
- GOSS, Gilroy & Associates Ltd (1989): Evaluation of the Job Development Program: Final Report. Program Evaluation Branch, Employment and Immigration Canada. Ottawa
- Gueron, J. M. (1991): Evaluation and Policy: The United States Experience with Employment and Training Programmes. In: OECD (Hrsg.), *Evaluating Labour Market and Social Programmes – The State of a Complex Art*. Paris, S. 167-173
- Hasan, A. (1991): Evaluation of Employment, Training and Social Programmes: An Overview of Issues. In: OECD (Hrsg.), *Evaluating Labour Market and Social Programmes – The State of a Complex Art*. Paris, S. 7-18
- Haveman, R. H./Saks, D. H. (1985): Transatlantic Lessons for Employment and Training Policy. In: *Industrial Relations*, Vol. 24, S. 20-36
- v. Hayek, F. A. (1976): Die Verwertung des Wissens in der Gesellschaft. In: *Individualismus und Wirtschaftliche Ordnung*. 2. erweiterte Auflage. Salzburg: Neugebauer, S. 103-121
- Johnson, G. E./Tomola J. D. (1977): The Fiscal Substitution Effect of Alternative Approaches to Public Service Employment Policy. In: *The Journal of Human Resources*, Vol. 12, S. 3-26
- Johnson, B./Wise, L. R. (1989): Getting Young People to Work? An Evaluation of Swedish Use Employment Policy. In: *International Labour Review*, Vol. 128, S. 337-356
- Long, D. A./Mallar, C. D./Thornton, C. V. D. (1981): Evaluating the Benefits and Costs of the Job Corps. In: *Journal of Policy Analysis and Management*, Vol. 1, S. 55-76
- Nathan, R. (1991): Evaluation Strategies with Particular Emphasis on Demonstration Studies. In: OECD (Hrsg.), *Evaluating Labour Market and Social Programmes – The State of a Complex Art*. Paris, S. 21-42
- Riddell, C. (1991): Evaluation of Manpower and Training Programs: The North American Experience. In: OECD (Hrsg.), *Evaluating Labour Market and Social Programmes – The State of a Complex Art*. Paris, S. 43-72
- Scharpf, F. W. (1982): Die Implementation des Sonderprogrammes im Vergleich. In: F. W. Scharpf/D. Garlichs/F. Maier/H. Maier, *Implementationsprobleme offensiver Arbeitsmarktpolitik. Das Sonderprogramm der Bundesregierung für Regionen mit besonderen Beschäftigungsproblemen*. Frankfurt, New York: Campus, S. 271-309
- Schellhaaß, H. M. (1991): Evaluation Strategies and Methods with Regard to Labour Market Programmes: A German Perspective. In: OECD (Hrsg.), *Evaluating Labour Market and Social Programmes – The State of a Complex Art*. Paris, S. 89-106
- Wilensky, H. L. (1985): Nothing Fails like Success: The Evaluation-Research Industry and Labor Market Policy. In: *Industrial Relations*, Vol. 24, S. 1-19
- Wiseman, M. (1976): Public Employment as Fiscal Policy. In: *Brookings Papers on Economic Activity*, Heft 1, S. 67-104
- Woodbury, S. A./Spiegelman, R. G. (1987): Bonuses to Workers and Employers to Reduce Unemployment: Randomized Trials in Illinois. In: *The American Economic Review*, Vol. 77, S. 513-530